# REPORT DOCUMENTATION



| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | March 1993 | Professional Paper |

**4. TITLE AND SUBTITLE**
ASSESSMENT OF GRAPHICS AND TEXT FORMATS FOR SYSTEM STATUS DISPLAYS

**5. FUNDING NUMBERS**
PR: CDB8
PE: 0602234N
WU: DN309119

**6. AUTHOR(S)**
W. A. Nugent and J. W. Broyles

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152-5001

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152-5001

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**



**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution is unlimited.

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

This study compared the relative effectiveness of three computer-based formats for displaying Navy system status data. Response speed and accuracy data were collected for each format on four tasks typically performed in a shipboard Combat Information Center (CIC). The three presentation formats were character readout (CRO), text-only, and text-graphics. Results showed the text-only and text-graphics formats produced faster, more accurate performance than the CRO on count and compare tasks; however, no reliable performance differences were found between presentation formats for identity and criterion tasks. Predictions concerning an advantage for the text-graphics format over the text-only format on certain types of tasks were not supported by the study findings. The practical applications and design implications of these findings are discussed.



Published in *Proceedings of the Human Factors Society 36th Annual Meeting,* Vol. 2, Oct 1992, pp 1464-1468.

**14. SUBJECT TERMS**
control and displays; man-machine interface and ASW; operator functions; command and control; decision support; decision aids; teleoperation; decision making; supervisory control

**15. NUMBER OF PAGES**

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | SAME AS REPORT |

| 21a. NAME OF RESPONSIBLE INDIVIDUAL | 21b. TELEPHONE (Include Area Code) | 21c. OFFICE SYMBOL |
|---|---|---|
| W. A. Nugent | (619) 553-8005 | Code 411 |

PROCEEDINGS of the HUMAN FACTORS SOCIETY 36th ANNUAL MEETING—1992



# ASSESSMENT OF GRAPHICS AND TEXT FORMATS FOR SYSTEM STATUS DISPLAYS

William A. Nugent and James W. Broyles

Naval Command, Control and Ocean Surveillance Center
Research, Development, Test and Evaluation Division
San Diego, CA 92152-5000

This study compared the relative effectiveness of three computer-based formats for displaying Navy system status data. Response speed and accuracy data were collected for each format on four tasks typically performed in a shipboard Combat Information Center (CIC). The three presentation formats were character readout (CRO), text-only, and text-graphics. Results showed the text-only and text-graphics formats produced faster, more accurate performance than the CRO on count and compare tasks; however, no reliable performance differences were found between presentation formats for identify and criterion tasks. Predictions concerning an advantage for the text-graphics format over the text-only format on certain types of tasks were not supported by the study findings. The practical applications and design implications of these findings are discussed.

## INTRODUCTION

Information overload can quickly overwhelm users of current Navy displays. Too often users must perform tedious search tasks and mental operations when processing displayed information; factors which may reduce the effectiveness of the combat system. New display technologies such as color and graphics offer the potential for decreasing the perceptual and cognitive burden on users of current status displays. Conversely, these technologies can also contribute to information overload if improperly used. There is a need to determine the most effective use of new display technologies if we are to contribute to better user-system performance.

The objective of this study was to compare the relative effectiveness of two experimental formats for presenting missile status data to a current format of displaying such data aboard U.S. Navy ships.

## METHOD

### Participants

Sixteen male U.S. Navy personnel (15 enlisted and one officer) participated in the study. The participants had served an average of 9.5 years in the Navy and had an average of 52 months Combat Information Center (CIC) experience. They ranged in age from 25 to 38 years, with a mean of 31.5 years and standard deviation of 4.6 years. All participants reported having normal or corrected-to-normal vision.

### Apparatus

A Macintosh Model IIci computer was used to present the task materials and to record test participants' responses and reaction times. Task materials were displayed on a 19-in. diagonal high-resolution color monitor. A numeric keypad with special keys configured for yes and no responses served as the input device. The monitor and keypad were installed in a mock-up of a prototype CIC operator console described in SPAWAR (1991).

### Display Materials

Three missile status data presentation formats were tested: character readout (CRO), text-only, and text-graphics. The CRO format (see Figure 1) was an exact replica of a current CIC display format. In addition to changes in color and layout, the experimental formats (see Figure 2) provide the following design changes to the current displays:

1. Missile inventory data is segregated for A- and B-rails on the FWD (forward) and AFT launchers, and
2. Row and column totals for different missile and launchers are displayed.

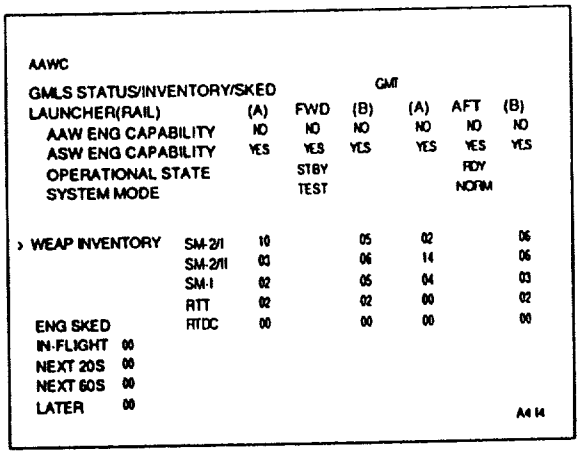



Figure 1. Example of the CRO format.

Further, launcher and system status fields for the experimental formats (shown in the lower panels of Figure 2) were reformatted and separated from missile inventory data. These displays also make use of fewer, more consistent terms and symbols to represent operational status data than the CRO format. (Figure 3 lists all symbols used in the experimental formats).

Besides differences in displaying launcher and system status data, note the bar charts to augment numeric data entries in Figure 2b. Their inclusion stemmed from the following Smith & Mosier (1986) guidelines:

2.4.2 When users must quickly scan and compare related sets of data, consider a graphic format to display the data. (p. 130)

2.4.9 When precise reading of a graphic display is required, annotate the display with actual data values to supplement their graphic representation. (p. 132)

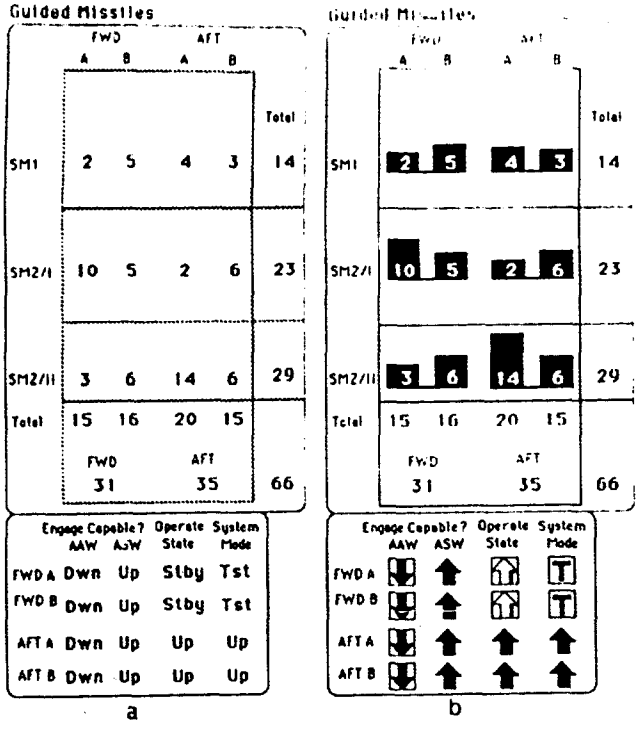



Figure 2. Example of the text-only format (a) and the text-graphics format (b)

Another difference between presentation formats concerned the use of color. All text and numeric characters in the CRO format appeared in phosphorescent green against a uniform black background. For the other formats, values of zero in the missile inventory fields appeared in red, as did symbols representing "Down" in the operational status fields. Symbols for the term "Up & On" appeared in green; those for the term "Degraded" in orange; and the background for the "Test" and "Notest" graphic icons in yellow. All other symbols were displayed as shown in Figures 2 & 3 with the exception of the background, which was a uniform light gray color.

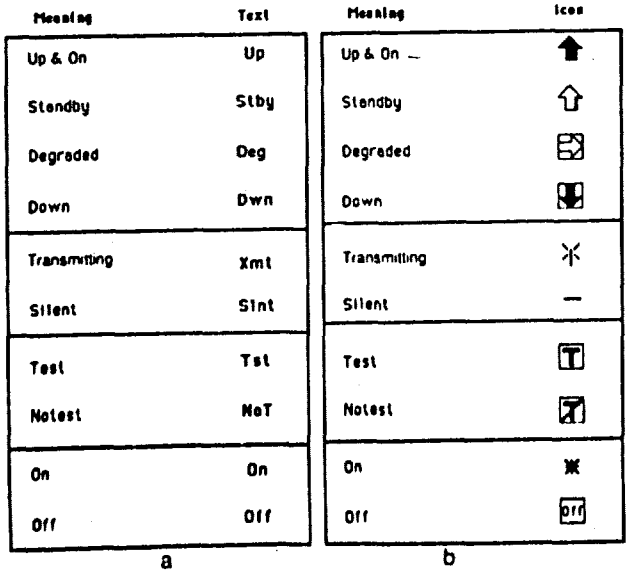



Figure 3. Symbols used in the text-only format (a) and text-graphics format (b)

## Test Materials

Test materials consisted of two parallel versions of 33 questions each. The questions, modeled after the work of Hitt et al. (1961), were divided into four categories of display reading tasks: count, compare, identify, and criterion. Examples of the two versions of test questions administered in each of the four task categories are listed in Table 1. The order of question administration was randomized within each test version and presentation format, yeilding six unique question orderings.

Table 1. Examples of test questions administered in each of four task categories.

| Task Category | No. of Items | Test Version | Example Questions |
|---|---|---|---|
| Count | 7 | 1 | What's the total standard missile inventory AFT? |
| | | 2 | What's the total standard missile inventory FWD? |
| Compare | 11 | 1 | Are there more standard missiles AFT A-rail than AFT B-rail? |
| | | 2 | Are there more standard missiles FWD A- rail than FWD B-rail? |
| Identify | 6 | 1 | What's the SM2/I missile inventory in the AFT A-rail? |
| | | 2 | What's the SM2/I missile inventory in the AFT B-rail? |
| Criterion | 9 | 1 | Are FWD & AFT launchers ASW engagement capable? |
| | | 2 | Are FWD & AFT launchers AAW engagement capable? |

## Procedure

Each participant was tested individually in two consecutive sessions in a classroom setting. At the outset of the session, instructions were read aloud to provide the participant a general orientation to the study, after which he completed a voluntary consent and background information form. On-screen instructions and examples of the stimulus materials to be used with the *first* presentation format were then shown. For the text-only and text-graphics formats, this included a 10-item task in which the participant associated text or graphic symbols with their proper operational terms. On-screen feedback was provided for incorrect responses to the association task. In addition, incorrectly answered items were reinserted (at random) into the queue until the participant met the task criterion of 100 percent response accuracy.

Participants then performed a 4-item practice exercise to become familiar with the presentation format to be used and actions to be performed. On-screen feedback was provided for incorrect responses to the practice exercises until the criterion of 100 percent response accuracy was met. Next, the 33 questions contained in the first presentation format were administered, followed by a short rest break. Identical procedures were used when administering the next set of questions in the alternative format. After completing the second session, participants evaluated the adequacy of the text or graphic symbols used in the experimental formats, then completed a separate user preference questionnaire.

## Experimental Hypotheses and Design

The following hypotheses served as the framework for comparing task performance as a function of display presentation format:

I. Participants will demonstrate faster response times and higher accuracy on all categories of tasks when using the experimental formats than when using the CRO format.

II. Since the display formats and information-processing requirements for count and identify tasks are quite similar, no performance differences will be found between the text-only and text-graphics formats on these tasks.

III. Because graphics allow a "quick scan" of the display to obtain initial estimates of high-low values, participants using the text-graphics format will perform compare tasks faster than participants using the text-only format.

IV. Because graphic icons convey the same information as text in a less "perceptually-cluttered" format, participants using the text-graphics format will perform criterion tasks faster and more accurately than participants using the text-only format.

A mixed factor design served as the model for data analysis. The between-subjects factor was assignment of test participants to one of the two experimental formats. The within-subjects factor was replications of each task category within alternative presentation formats: CRO vs. text-only and CRO vs. text-graphics. Test version and order of administering the two presentation formats was counterbalanced using Latin squares. Dependent measures were response speed and accuracy. Response speed was measured from the initial presentation of a question to an entry response. Accuracy was the percentage of questions answered correctly by the test participant.

## RESULTS

### Preliminary Analyses

Two separate analysis of variance (ANOVAs) were conducted preparatory to the main tests of the experimental hypotheses. The first tested for differences in performance speed and accuracy as a function of test version and display format presentation order; the second as a function of test participant age, months of CIC experience, and computer proficiency.

Results showed nonsignificant F-ratios for all main effects and interactions. Accordingly, it was concluded: (1) the two question sets are equivalent, (2) order of administering the presentation formats did not produce any asymmetrical transfer effects, and (3) the participants' background characteristics did not have any differential effects on task performance speed or accuracy.

### Main Analyses

Hypothesis I. Because the four categories of tasks contained unequal numbers of questions, relevant data for this hypothesis were analyzed in 8 mixed factor ANOVAs—one for each task category separately on the speed and accuracy measures. Results, summarized in Tables 2 & 3, show the experimental formats were superior to the CRO in terms of response accuracy and speed for the categories of count and compare tasks.

Specifically, examination of Table 2 shows when using the experimental formats, participants had an *absolute* difference of 14 and 11 percent higher accuracy on count and compare tasks than when using the CRO format. This translates into an 18 and 13 percent *relative* improvement over the CRO for these tasks. Similarly, Table 3 shows a speed advantage for the experimental formats over the CRO: participants took 54 and 38 percent less time when responding to count and compare items, respectively.

Table 2 Summary of analysis of variance and means for response accuracy.

| | | | Mean Accuracy (in percent) | |
|---|---|---|---|---|
| Category | df | F | Exper. | CRO |
| Count | 1, 110 | 9.02** | 92 | 78 |
| Compare | 1, 174 | 8.40** | 93 | 82 |
| Identify | 1, 94 | 0.22 | 91 | 89 |
| Criterion | 1, 142 | 0.40 | 96 | 97 |

**Significant at $p<.01$

Table 3. Summary of analysis of variance and means for response speed.

| | | | Mean Speed (in seconds) | |
|---|---|---|---|---|
| Category | df | F | Exper. | CRO |
| Count | 1, 110 | 18.31** | 7.64 | 11.79 |
| Compare | 1, 174 | 38.90** | 11.03 | 15.18 |
| Identify | 1, 94 | 0.27 | 8.04 | 8.23 |
| Criterion | 1, 142 | 0.0004 | 8.41 | 8.42 |

**Significant at $p<.01$

Contrary to the Hypothesis I predictions, no reliable differences were found between the CRO and experimental formats relative to performance speed or accuracy on identify and criterion tasks.

Hypothesis II. This hypothesis predicted no performance differences between the experimental formats for count and identify tasks. Relevant data for this hypothesis were extracted from the mixed factor ANOVAs described above. As predicted, no significant differences were found between the text-only and text-graphics formats relative to performance speed or accuracy on count and identify tasks. The means tested in these and other comparisons are presented in Figures 4 & 5.

Hypothesis III. This hypothesis predicted faster response times for the text-graphics format than the text-only format on compare tasks. ANOVA results showed no reliable difference between these presentation formats in terms of performance speed ($F(1,174) = 0.52$, $p<.48$); thus failing to support Hypothesis III predictions.

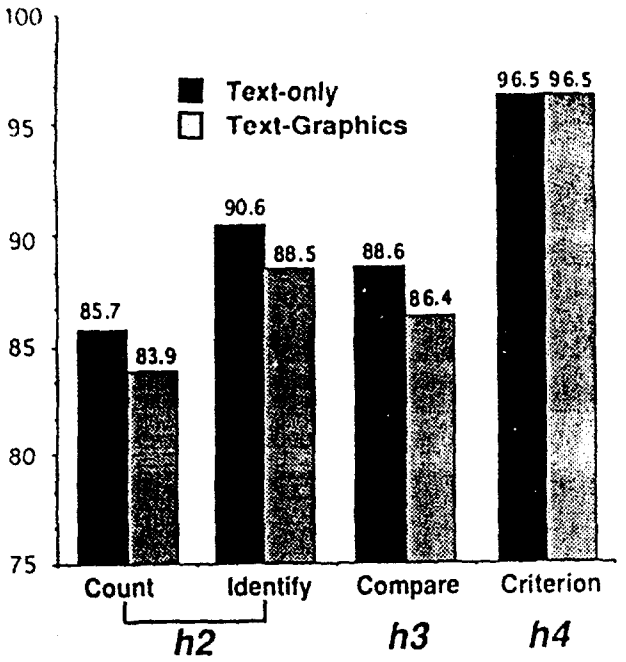
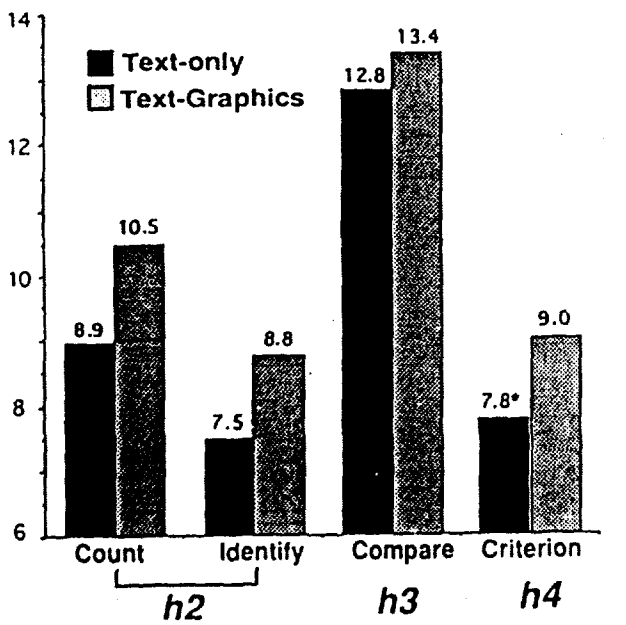



Figure 4. Response accuracy for two experimental presentation formats on four tasks.




* Significant at $p < .01$

Figure 5. Response speed for two experimental presentation formats on four tasks.

Hypothesis IV. This hypothesis predicted faster, more accurate performance for the text-graphics format than the text-only format on criterion tasks. While no difference was found between these formats relative to accuracy, the main effect for performance speed did reach statistical significance: $F(1,142) = 5.50$, $p<.02$. Contrary to hypothesis predictions, participants who used the text-only format had faster response times (7.82 seconds) than participants who used the text-graphics format (9.01 seconds).

User Preferences

Participant background characteristics and experimental presentation formats were also included as factors in analyses of ratings on two user preference surveys. In the first survey, participants rated how well each text or graphic symbol (see Figure 3) represented its corresponding operational term. Scale values for this survey ranged from 1—not representative to 5—highly representative. In the second survey, participants rated items concerning their attitudes toward the experimental formats on a different 5-point scale.

No reliable differences were found in ratings on either survey as a function of participant background or presentation format factors. Significant differences were found, however, between ratings of text and graphic symbols on the first survey. Specifically, average ratings for the "Silent" symbol were lower than all other symbols except "Notest" and "Transmitting" at $p < .05$.

Average ratings assigned to the experimental formats on the second survey were generally quite favorable (see Figure 6). Participants rated these formats as "Somewhat better" than current CIC displays (Mean 4.2). Positive ratings were also assigned to items concerning the participants' overall attitude toward the experimental formats (Mean 4.5), how easy they were to learn (Mean 4.8), and their ease of use (Mean 4.1).

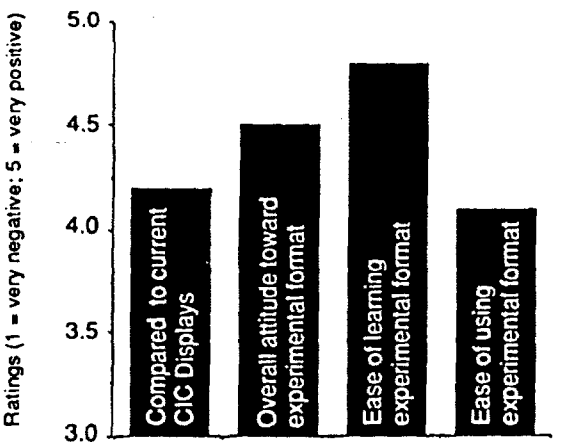



Figure 6. Average user preference ratings for experimental presentation formats.

## CONCLUSIONS AND DESIGN IMPLICATIONS

While results show the experimental formats offered both a speed and accuracy advantage over the CRO for count and compare tasks, no reliable differences were found between these formats for identify and criterion tasks.

Several factors may explain these findings. First, the layout of the experimental formats clearly reduced the time needed to locate and process pertinent data—particularly for items involving row and column totals. These layouts were of little benefit for identify and criterion tasks. The majority of these tasks involved locating discrete values within a single row or column.

Second, the CRO format may have offered better stimulus-response compatibility for certain criterion task items than the experimental formats. In several cases, YES and NO terms used to describe the operational status of missile

launchers in the CRO format were synonymous with correct responses to test items. The experimental formats, in contrast, required text or graphic symbols to be translated into operational terms prior to formulating a yes or no response.

Conflicting results obtained for comparisons between the experimental formats used in this study also warrant discussion. Specifically, the lack of a speed advantage for the text-graphics format on compare tasks (Hypothesis III), may be explained in two ways. First, it is possible that participants simply ignored the relative height of the bar-graphs when making comparisons; instead, focusing on the displayed numeric values. Second, the wording of these items may have effectively precluded a quick scan of the display to obtain estimates of high-low values. Perhaps a more appropriate way of testing for a speed advantage in the text-graphics format would involve rewording the items (e.g., Are there enough SM2-II missiles FWD to engage threat X?). Here, the comparison is implicit—tapping internalized rules concerning the number of missiles required to engage the threat. When displayed in the text-graphics format, bar-graphs may then function as a type of "fuel gauge," allowing the operator to tell at a glance if sufficient missiles were available.

Finally, the speed advantage for the text-only format on criterion tasks (Hypothesis IV), may be attributed to added information-processing "overhead" for the text-graphics format. That is, participants who used the text-only format may have spent slightly less time translating displayed symbols into their operationally-equivalent terms than participants who used the text-graphics format.

Based on the study findings, a specific recommendation for improving the design of current Navy status displays is as follows:

- The text-only format, exemplified in Figure 2a, is the best choice for upgrading the display of missile and system status data from the current CRO format.

In addition, the following guidelines provide possible explanations of these results:

1. To facilitate rapid scanning and assimilation of multiple data elements in tabular form, segregate the display such that only the information which is deemed relevant to the task at hand is perceived and/or acted on by users.

2. Ensure that data displayed in tabular form is organized into clearly defined and labeled rows and columns.

3. If the tabular display is to be used for tasks other than locating discrete values within a particular row or column, provide a capability for generating and dynamically updating totals—by applicable rows, columns, or aggregations of rows and columns.

4. Ensure that terms or symbols representing operational status or state data are directly compatible, and preferably, synonymous with responses required of the operator when processing these data.

## ACKNOWLEDGMENTS

The authors express gratitude to Tim Warkentin for designing the computer software used to present the task materials and record test participants' responses and reaction times. Appreciation is also expressed to Susan Bemis for numerous hours of data reduction and analysis. Opinions expressed within are those of the authors, and do not represent official U.S. Navy or Department of Defense policy, position, or decision. This work was supported under the Office of Naval Technology Human Factors Project, and Naval Sea Systems Command PMS 400 Advanced AEGIS Human Factors Studies.

## REFERENCES

Hitt, W.D, H. Schultz, C. Christner, H. Ray & L. Coffey. Development of design criteria for intelligence display formats, *Human Factors*, 1961, 3, 86-148.

Smith, S.L & Mosier, J.N. Guidelines for designing user interface software, MTR 10090, MITRE Corporation, Bedford, MA, August 1986.

SPAWAR, September 1991. STANAG 4420 ratification recommendations: final report. Space and Naval Warfare Systems Command, Code 2311C, Washington, D.C.